

EXHIBIT B

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12Q 1/68, A01K 67/027, C12N 15/12		A1	(11) International Publication Number: WO 99/29896 (43) International Publication Date: 17 June 1999 (17.06.99)		
(21) International Application Number: PCT/CA98/01133		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).			
(22) International Filing Date: 7 December 1998 (07.12.98)					
(30) Priority Data: 2,218,199 9 December 1997 (09.12.97) CA					
(71) Applicant (for all designated States except US): McGILL UNIVERSITY [CA/CA]; 845 Sherbrooke Street West, Montréal, Québec H3A 2T5 (CA).					
(72) Inventors and (75) Inventors/Applicants (for US only): ROULEAU, Guy, A. [CA/CA]; Appartement 7, 4850 Côte Saint-Luc, Montréal, Québec H3W 2H2 (CA). BRAIS, Bernard [CA/CA]; 745 Champagneau, Outremont, Québec H2V 3P9 (CA).		Published <i>With international search report.</i>			
(74) Agents: COTE, France et al.; Swabey Ogilvy Renault, Suite 1600, 1981 McGill College Avenue, Montréal, Québec H3A 2Y3 (CA).					
(54) Title: SHORT GCG EXPANSIONS IN THE PAB II GENE FOR OCULOPHARYNGEAL MUSCULAR DYSTROPHY AND DIAGNOSTIC THEREOF					
(57) Abstract					
<p>The present invention relates to a human PAB II gene containing transcribed polymorphic GCG repeat, which comprises a sequence as set forth in SEQ ID NO:3, which includes introns and flanking genomic sequence. The allelic variants of GCG repeat of the human PAB II gene are associated with a disease related with protein accumulation in nucleus, such as polyalanine accumulation, a disease related with swallowing difficulties, such as oculopharyngeal muscular dystrophy. The present invention also relates to a method for the diagnosis of a disease with protein accumulation in nucleus, which comprises the steps of: a) obtaining a nucleic acid sample of said patient; and b) determining allelic variants of GCG repeat of the gene of claim 1, and wherein long allelic variants are indicative of a disease related with protein accumulation in nucleus.</p>					

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

SHORT GCG EXPANSIONS IN THE PAB II GENE FOR OCULO-
PHARYNGEAL MUSCULAR DYSTROPHY AND DIAGNOSTIC THEREOF

BACKGROUND OF THE INVENTION

5 (a) Field of the Invention

The invention relates to PAB II gene, and its uses thereof for the diagnosis, prognosis and treatment of a disease related with protein accumulation in nucleus, such as oculopharyngeal muscular dystrophy.

10 (b) Description of Prior Art

Autosomal dominant oculopharyngeal muscular dystrophy (OPMD) is an adult-onset disease with a worldwide distribution. It usually presents in the sixth decade with progressive swallowing difficulties (dysphagia), eye lid drooping (ptosis) and proximal limb weakness. Unique nuclear filament inclusions in skeletal muscle fibers are its pathological hallmark (Tome, F.M.S. & Fardeau, *Acta Neuropath.* **49**, 85-87 (1980)). We isolated the poly(A) binding protein II (PAB II) gene from a 217 kb candidate interval in chromosome 14q11. A (GCG)₆ repeat encoding a polyalanine tract located at the N-terminus of the protein was expanded to (GCG)₈₋₁₃ in the 144 OPMD families screened. More severe phenotypes were observed in compound heterozygotes for the (GCG)₉ mutation and a (GCG)₇ allele found in 2% of the population, whereas homozygosity for the (GCG)₇ allele leads to autosomal recessive OPMD. Thus the (GCG)₇ allele is an example of a polymorphism which can act as either a modifier of a dominant phenotype or as a recessive mutation. Pathological expansions of the polyalanine tract may cause mutated PAB II oligomers to accumulate as filament inclusions in nuclei.

It would be highly desirable to be provided with a tool for the diagnosis, prognosis and treatment of a disease related with polyalanine accumulation in nucleus, such as oculopharyngeal muscular dystrophy.

SUMMARY OF THE INVENTION

One aim of the present invention is to provide a tool for the diagnosis, prognosis and treatment of a disease related with polyalanine accumulation in nucleus, such as oculopharyngeal muscular dystrophy.

In accordance with the present invention there is provided a human PAB II gene containing transcribed polymorphic GCG repeat, which comprises a sequence as set forth in Fig. 4, which includes introns and flanking genomic sequence.

The allelic variants of GCG repeat of the human PAB II gene are associated with a disease related with protein accumulation in nucleus, such as polyalanine accumulation, or with a disease related with swallowing difficulties, such as oculopharyngeal muscular dystrophy.

In accordance with the present invention there is also provided a method for the diagnosis of a disease with protein accumulation in nucleus, which comprises the steps of:

- a) obtaining a nucleic acid sample of said patient; and
- b) determining allelic variants of GCG repeat of the gene of the human PAB II gene, and wherein long allelic variants are indicative of a disease related with protein accumulation in nucleus, such as polyalanine accumulation and oculopharyngeal muscular dystrophy.

The long allelic variants have from about 245 to about 263 bp in length.

In accordance with the present invention there is also provided a non-human mammal model for the PAB II gene of the human PAB II gene, whose germ cells and somatic cells are modified to express at least one allelic variant of the PAB II gene and wherein said

allelic variant of the PAB II being introduced into the mammal, or an ancestor of the mammal, at an embryonic stage.

5 In accordance with the present invention there is also provided a method for the screening of therapeutic agents for the prevention and/or treatment of oculopharyngeal muscular dystrophy, which comprises the steps of:

- 10 a) administering said therapeutic agents to the non-human mammal of the present invention or oculopharyngeal muscular dystrophy patients; and
- b) evaluating the prevention and/or treatment of development of oculopharyngeal muscular dystrophy in said mammal or said patients.

15 In accordance with the present invention there is also provided a method to identify genes part of or interacting with a biochemical pathway affected by PAB II gene, which comprises the steps of:

- 20 a) designing probes and/or primers using the hGT1 gene of the PAB II gene and screening oculopharyngeal muscular dystrophy patients samples with said probes and/or primers; and
- b) evaluating the identified gene role in oculopharyngeal muscular dystrophy patients.

25

BRIEF DESCRIPTION OF THE DRAWINGS

Figs. 1A-B illustrate the positional cloning of the PAB II gene;

30 Figs. 2A-G illustrate the OPMD (GCG)_n expansion sizes and sequence of mutations (SEQ ID NOS:1-2);

Fig. 3 illustrates the age distribution of swallowing time (st) for French Canadian OPMD carriers of the (GCG)₉ mutation; and

Fig. 4 illustrates the nucleotide sequence of human poly(A) binding protein II (hPAB II) (SEQ ID NO:3).

5 **DETAILED DESCRIPTION OF THE INVENTION**

In order to identify the gene mutated in OPMD, we constructed a 350 kb cosmid contig between flanking markers D14S990 and D14S1457 (Fig. 1A). Positions of the PAB II selected cDNA clones in relation to the 10 EcoRI restriction map and the Genealogy-based Estimate of Historical Meiosis (GEHM)-derived candidate interval (Rommens, J.M. et al., in *Proceedings of the third international workshop on the identification of transcribed sequences* (eds. Hochgeschwender, U. & Gardiner, K.) 65-79 (Plenum, New York, 1994)).

The human poly(A) binding protein II gene (PAB II) is encoded by the nucleotide sequence as set forth in Fig. 4.

Twenty-five cDNAs were isolated by cDNA selection from the candidate interval (Rommens, J.M. et al., in *Proceedings of the third international workshop on the identification of transcribed sequences* (eds. Hochgeschwender, U. & Gardiner, K.) 65-79 (Plenum, New York, 1994)). Three of these hybridized to a common 20 kb EcoRI restriction fragment and showed high sequence homology to the bovine poly(A) binding protein II gene (bPAB II) (Fig. 1A). The PAB II gene appeared to be a good candidate for OPMD because it mapped to the genetically defined 0.26 cM candidate interval in 14q11 (Fig. 1A), its mRNA showed a high level of expression in skeletal muscle, and the PAB II protein is exclusively localized to the nucleus (Krause, S. et al., *Exp. Cell Res.* **214**, 75-82 (1994)) where it acts as a factor in mRNA polyadenylation (Whale, E., *Cell* **66**, 35 759-768 (1991); Whale, E. et al., *J. Biol. Chem.* **268**,

2937-2945 (1993); Bienroth, S. et al., *EMBO J.* **12**, 585-594 (1993)).

We subcloned a 8 kb HindIII genomic fragment containing the PAB II gene, and sequenced 6002 bp (GenBank: AF026029) (Nemeth, A. et al., *Nucleic Acids Res.* **23**, 4034-4041 (1995)) (Fig. 1B). Genomic structure of the PAB II gene, and position of the OPMD (GCG)_n expansions. Exons are numbered. Introns 1 and 6 are variably present in 60% of cDNA clones. ORF, open reading frame; cen, centromere and tel, telomere.

The coding sequence was based on the previously published bovine sequence (GenBank: X89969) and the sequence of 31 human cDNAs and ESTs. The gene is composed of 7 exons and is transcribed in the cen-*q*ter orientation (Fig. 1B). Multiple splice variants are found in ESTs and on Northern blots (Nemeth, A. et al., *Nucleic Acids Res.* **23**, 4034-4041 (1995)). In particular, introns 1 and 6 are present in more than 60% of clones (Fig. 1B) (Nemeth, A. et al., *Nucleic Acids Res.* **23**, 4034-4041 (1995)). The coding and protein sequences are highly conserved between human, bovine and mouse (GenBank: U93050). 93% of the PAB II sequence was readily amenable to RT-PCR- or genomic-SSCP screening. No mutations were uncovered using both techniques. However, a 400 bp region of exon 1 containing the start codon could not be readily amplified. This region is 80% GC rich. It includes a (GCG)₆ repeat which codes for the first six alanines of a homopolymeric stretch of 10 (Fig. 2G). Nucleotide sequence of the mutated region of PAB II. Amino acid sequences of the N-terminus polyalanine stretch and position of the OPMD alanine insertions.

Special conditions were designed to amplify by PCR a 242 bp genomic fragment including this GCG-repeat. The (GCG)₆ allele was found in 98% of French

Canadian non-OPMD control chromosomes, whereas 2% of chromosomes carried a (GCG)₇ polymorphism (n=86) (Brais, B. et al., *Hum. Mol. Genet.* **4**, 429-434 (1995)).

Screening OPMD cases belonging to 144 families showed in all cases a PCR product larger by 6 to 21 bp than that found in controls (Fig. 2A). (GCG)₆ normal allele (N) and the six different (GCG)_n expansions observed in 144 families.

Sequencing of these fragments revealed that the increased sizes were due to expansions of the GCG repeat (Fig. 2G). Fig. 2F shows the sequence of the (GCG)₉ French Canadian expansion in a heterozygous parent and his homozygous child. Partial sequence of exon 1 in a normal (GCG)₆ control (N), a heterozygote (ht.) and a homozygote (hm.) for the (GCG)₉-repeat mutation. The number of families sharing the different (GCG)_n-repeats expansions is shown in Table 1.

Table 1

Number of families sharing the different dominant (GCG)_n OPMD mutations

Mutations	Polyalanine	Families
(GCG) ₈	12	4
(GCG) ₉	13	99
(GCG) ₁₀	14	19
(GCG) ₁₁	15	16
(GCG) ₁₂	16	5
(GCG) ₁₃	17	1
Total		144

†, 10 alanine residues in normal PAB II.

The (GCG)₉ expansion shared by 70 French Canadian families is the most frequent mutation we observed (Table 1). The (GCG)₉ expansion is quite stable, with a single doubling observed in family F151 in an estimated 598 French Canadian meioses (Fig. 2C). The doubling of

the French Canadian (GCG)₉ expansion is demonstrated in Family F151.

This contrasts with the unstable nature of previously described disease-causing triplet-repeats (Rosenberg, R.N., *New Eng. J. Med.* **335**, 1222-1224 (1996)).

Genotyping of all the participants in our clinical study of French Canadian OPMD provided molecular insights into the clinical variability observed in this condition. The genotypes for both copies of the PAB II mutated region were added to an anonymous version of our clinical database of 176 (GCG)₉ mutation carriers (Brais, B. et al., *Hum. Mol. Genet.* **4**, 429-434 (1995)). Severity of the phenotype can be assessed by the swallowing time (st) in seconds taken to drink 80 cc of ice-cold water (Brais, B. et al., *Hum. Mol. Genet.* **4**, 429-434 (1995); Bouchard, J.-P. et al., *Can. J. Neurol. Sci.* **19**, 296-297 (1992)). The late onset and progressive nature of the muscular dystrophy is clearly illustrated in heterozygous carriers of the (GCG)₉ mutation (bold curve in Fig. 3) when compared the average st of control (GCG)₆ homozygous participants (n=76, thinner line in Fig. 3). The bold curve represents the average OPMD st for carriers of only one copy of the (GCG)₉ mutation (n=169), while the thinner line corresponds to the average st for (GCG)₆ homozygous normal controls (n=76). The black dot corresponds to the st value for individual VIII. Roman numerals refer to individual cases shown in Figs. 2B, 2D and discussed in the text. Genotype of a homozygous (GCG)₉ case and her parents (Fig. 2B). Independent segregation of the (GCG)₇ allele. Case V has a more severe OPMD phenotype (Fig. 2D).

Two groups of genotypically distinct OPMD cases have more severe swallowing difficulties. Individuals

I, II, and III have an early-onset disease and are homozygous for the (GCG)₉ expansion ($P < 10^{-5}$) (Figs. 2B, F). Cases IV, V, VI and VII have more severe phenotypes and are compound heterozygotes for the 5 (GCG)₉ mutation and the (GCG)₇ polymorphism ($P < 10^{-5}$). In Fig. 2D the independent segregation of the two alleles is shown. Case V, who inherited the French Canadian (GCG)₉ mutation and the (GCG)₇ polymorphism, is more symptomatic than his brother VIII who carries 10 the (GCG)₉ mutation and a normal (GCG)₆ allele (Figs. 2D and 3). The (GCG)₇ polymorphism thus appears to be a modifier of severity of dominant OPMD. Furthermore, the (GCG)₇ allele can act as a recessive mutation. This was documented in the French patient IX who 15 inherited two copies of the (GCG)₇ polymorphism and has a late-onset autosomal recessive form of OPMD (Fig. 2E). Case IX, who has a recessive form of OPMD, is shown to have inherited two copies of the (GCG)₇ polymorphism.

20 This is the first description of short trinucleotide repeat expansions causing a human disease. The addition of only two GCG repeats is sufficient to cause dominant OPMD. OPMD expansions do not share the cardinal features of "dynamic mutations". The GCG expansions 25 are not only short they are also meiotically quite stable. Furthermore, there is a clear cut-off between the normal and abnormal alleles, a single GCG expansion causing a recessive phenotype. The PAB II (GCG)₇ allele is the first example of a relatively frequent allele 30 which can act as either a modifier of a dominant phenotype or as a recessive mutation. This dosage effect is reminiscent of the one observed in a homozygote for two dominant synpolydactyly mutations. In this case, the patient had more severe deformities because she inherited 35 two duplications causing an expansion in the

polyalanine tract of the HOXD13 protein (Akarsu, A.N. et al., *Hum. Mol. Genet.* **5**, 945-952 (1996)). A duplication causing a similar polyalanine expansion in the a subunit 1 gene of the core-binding transcription factor 5 (CBF α 1) has also been found to cause dominant cleidocranial dysplasia (Mundlos, S. et al., *Cell* **89**, 773-779 (1997)). The mutations in these two rare diseases are not triplet-repeats. They are duplications of "cryptic repeats" composed of mixed synonymous codons and are 10 thought to result from unequal crossing over (Warren, S.T., *Science* **275**, 408-409 (1997)). In the case of OPMD, slippage during replication causing a reiteration of the GCG codon is a more likely mechanism (Wells, D.R., *J. Biol. Chem.* **271**, 2875-2878 (1996)).

15 Different observations converge to suggest that a gain of function of PAB II may cause the accumulation of nuclear filaments observed in OPMD (Tome, F.M.S. & Fardeau, *Acta Neuropath.* **49**, 85-87 (1980)). PAB II is found mostly in dimeric and oligomeric form (Nemeth, A. et al., *Nucleic Acids Res.* **23**, 4034-4041 (1995)). It is possible that the polyalanine tract plays a role in polymerization. Polyalanine stretches have been found in many other nuclear proteins such as the HOX proteins, but their functions is still unknown (Davies, 20 S.W. et al., *Cell* **90**, 537-548 (1997)). Alanine is a highly hydrophobic amino acid present in the cores of proteins. In dragline spider silk, polyalanine stretches are thought to form B-sheet structures important in ensuring the fibers' strength (Simmons, A.H. et 25 al., *Science* **271**, 84-87 (1996)). Polyalanine oligomers have also been shown to be extremely resistant to chemical denaturation and enzymatic degradation (Forood, B. et al., *Bioch. and Biophys. Res. Com.* **211**, 7-13 (1995)). One can speculate that PAB II oligomers 30 comprised of a sufficient number of mutated molecules 35

might accumulate in the nuclei by forming undegradable polyalanine rich macromolecules. The rate of the accumulation would then depend on the ratio of mutated to non-mutated protein. The more severe phenotypes 5 observed in homozygotes for the (GCG)₉ mutations and compound heterozygotes for the (GCG)₉ mutation and (GCG)₇ allele may correspond to the fact that in these cases PAB II oligomers are composed only of mutated proteins. The ensuing faster filament accumulation 10 could cause accelerated cell death. The recent description of nuclear filament inclusions in Huntington's disease, raises the possibility that "nuclear toxicity" caused by the accumulation of mutated homopolymeric domains is involved in the molecular pathophysiology of 15 other triplet-repeat diseases (Davies, S.W. et al., *Cell* **90**, 537-548 (1997); Scherzinger, E. et al., *Cell* **90**, 549-558 (1997); DiFiglia, M. et al., *Science* **277**, 1990-1993 (1997)). Future immunocytochemical and 20 expression studies will be able to test this pathophysiological hypothesis and provide some insight into why certain muscle groups are more affected while all tissues express PAB II.

Methods

Contig and cDNA selection

25 The cosmid contig was constructed by standard cosmid walking techniques using a gridded chromosome 14-specific cosmid library (Evans, G.A. et al., *Gene* **79**, 9-20 (1989)). The cDNA clones were isolated by cDNA 30 selection as previously described (Rommens, J.M. et al., in *Proceedings of the third international workshop on the identification of transcribed sequences* (eds. Hochgeschwender, U. & Gardiner, K.) 65-79 (Plenum, New York, 1994)).

Cloning of the PAB II gene. Three cDNA clones 35 corresponding to PAB II were sequenced (Sequenase,

USB). Clones were verified to map to cosmids by Southern hybridization. The 8 kb HindIII restriction fragment was subcloned from cosmid 166G8 into pBluescriptII (SK) (Stratagene). The clone was sequenced using primers derived from the bPABII gene and human EST sequences. Sequencing of the PAB II introns was done by primer walking.

PAB II mutation screening and sequencing. All cases were diagnosed as having OPMD on clinical grounds (Brais, B. et al., *Hum. Mol. Genet.* **4**, 429-434 (1995)). RT-PCR- and genomic SSCP analyses were done using standard protocols (Lafrenière, R.G. et al., *Nat. Genet.* **15**, 298-302 (1997)). The primers used to amplify the PAB II mutated region were: 5'-CGCAGTGCCCCGCCTTAGA-3' (SEQ ID NO:4) and 5'-ACAAGATGGCGCCGCCGCCGC-3' (SEQ ID NO:5). PCR reactions were performed in a total volume of 15 ml containing: 40 ng of genomic DNA; 1.5 mg of BSA; 1 mM of each primer; 250 mM dCTP and dTTP; 25 mM dATP; 125 mM of dGTP and 125 mM of 7-deaza-dGTP (Pharmacia); 7.5% DMSO; 3.75 mCi[35S]dATP, 1.5 unit of Taq DNA polymerase and 1.5 mM MgCl₂ (Perkin Elmer). For non-radioactive PCR reactions the [35S]dATP was replaced by 225 mM of dATP. The amplification procedure consisted of an initial denaturation step at 95°C for five minutes, followed by 35 cycles of denaturation at 95°C for 15 s, annealing at 70°C for 30 s, elongation at 74°C for 30 s and a final elongation at 74°C for 7 min. Samples were loaded on 5% polyacrylamide denaturing gels. Following electrophoresis, gels were dried and autoradiographs were obtained. Sizes of the inserts were determined by comparing to a standard M13 sequence (Sequenase, USB). Fragments used for sequencing were gel-purified. Sequencing of the mutated fragment using the Amplicycle kit (Perkin Elmer) was done with the 5'-

CGCAGTCCCCGCCTAGAGGTG-3' (SEQ ID NO:6) primer at an elongation temperature of 68°C.

Stability of (GCG)-repeat expansions. The meiotic stability of the (GCG)9-repeat was estimated based on our large French Canadian OPMD cohort. We previously established that a single ancestral OPMD carrier chromosome was introduced in the French Canadian population by three sisters in 1648. Seventy of the seventy one French Canadian OPMD families tested to date segregate a (GCG)9 expansion. However, in family F151, the affected brother and sister, despite sharing the French Canadian ancestral haplotype, carry a (GCG)12 expansion twice the size of the ancestral (GCG)9 mutation (Fig. 2C). In our founder effect study, we estimated that 450 (304-594) historical meioses shaped the 123 OPMD cases belonging to 42 of the 71 enrolled families. Our screening of our full set of participants allowed us to identify another 148 (GCG)9 carrier chromosomes. Therefore, we estimate that a single mutation of the (GCG)9 expansion has occurred in 598 (452-742) meioses.

Genotype-phenotype correlations. 176 carriers of at least one copy of the (GCG)9 mutation were examined during the early stage of the linkage study. All were asked to swallow 80 cc of ice-cold water as rapidly as possible. Testing was stopped after 60 seconds. The swallowing time (st) was validated as a sensitive test to identify OPMD cases (Brais, B. et al., *Hum. Mol. Genet.* **4**, 429-434 (1995); Bouchard, J.-P. et al., *Can. J. Neurol. Sci.* **19**, 296-297 (1992)). The st values for 76 (GCG)6 homozygotes normal controls is illustrated in Fig. 3. Analyses of variance were computed by two-way ANOVA (SYSTAT package). For the (GCG)9 homozygotes their mean st value was compared to the mean value for all (GCG)9 heterozygotes aged 35-40 ($P < 10^{-5}$). For the (GCG)9 and (GCG)7 compound heterozygotes their mean st

value was compared to the mean value for all (GCG) 9 heterozygotes aged 45-65 ($P < 10^{-5}$).

While the invention has been described in connection with specific embodiments thereof, it will be understood that it is capable of further modifications and this application is intended to cover any variations, uses, or adaptations of the invention following, in general, the principles of the invention and including such departures from the present disclosure as come within known or customary practice within the art to which the invention pertains and as may be applied to the essential features hereinbefore set forth, and as follows in the scope of the appended claims.

WHAT IS CLAIMED IS:

1. A human PAB II gene containing transcribed polymorphic GCG repeat, which comprises a sequence as set forth in SEQ ID NO:3, which includes introns and flanking genomic sequence.
2. The gene of claim 1, wherein allelic variants of GCG repeat are associated with a disease related with protein accumulation in nucleus.
3. The gene of claim 2, wherein said protein accumulation is polyalanine accumulation.
4. The gene of claim 1, wherein allelic variants of GCG repeat are associated with a disease related with swallowing difficulties.
5. The gene of claim 1, wherein said disease is oculopharyngeal muscular dystrophy.
6. A method for the diagnosis of a disease with protein accumulation in nucleus, which comprises the steps of:
 - a) obtaining a nucleic acid sample of said patient; and
 - b) determining allelic variants of GCG repeat of the gene of claim 1, and wherein long allelic variants are indicative of a disease related with protein accumulation in nucleus.
7. The method of claim 6, wherein said disease is oculopharyngeal muscular dystrophy.

8. The method of claim 7, wherein said long allelic variants have from about 245 to about 263 bp in length.

9. A non-human mammal model for the PAB II gene of claim 1, whose germ cells and somatic cells are modified to express at least one allelic variant of the PAB II gene and wherein said allelic variant of the PAB II being introduced into the mammal, or an ancestor of the mammal, at an embryonic stage.

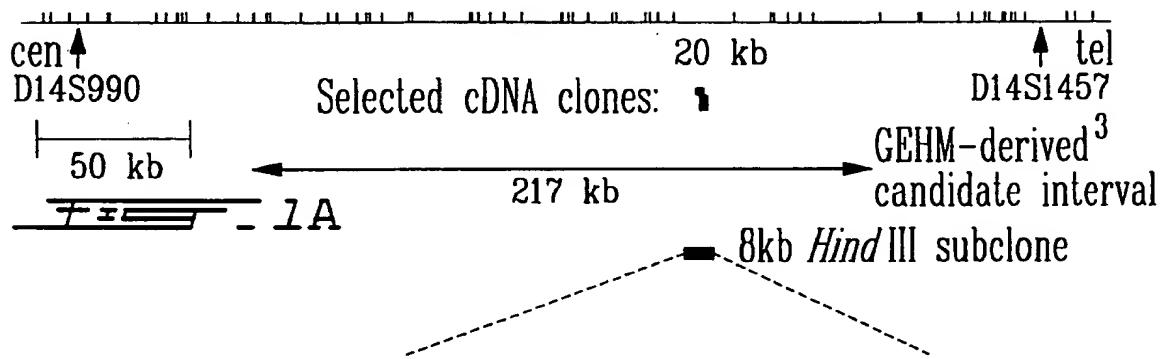
10. A method for the screening of therapeutic agents for the prevention and/or treatment of oculopharyngeal muscular dystrophy, which comprises the steps of:

- a) administering said therapeutic agents to the non-human mammal of claim 9 or oculopharyngeal muscular dystrophy patients; and
- b) evaluating the prevention and/or treatment of development of oculopharyngeal muscular dystrophy in said mammal or said patients.

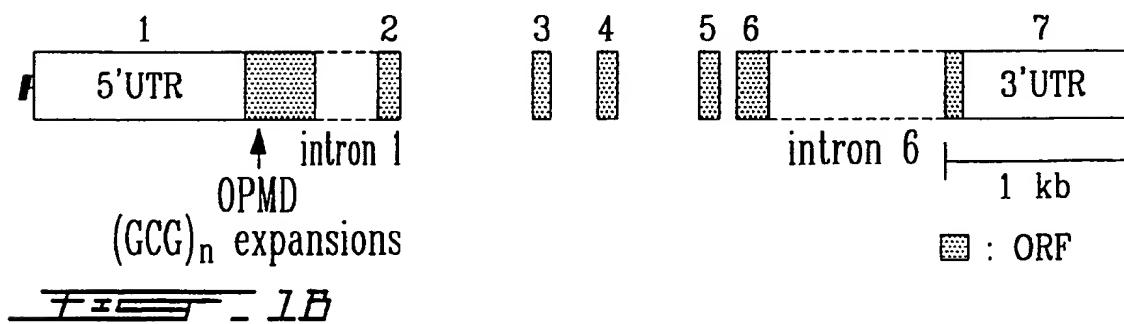
11. A method to identify genes part of or interacting with a biochemical pathway affected by PAB II gene, which comprises the steps of:

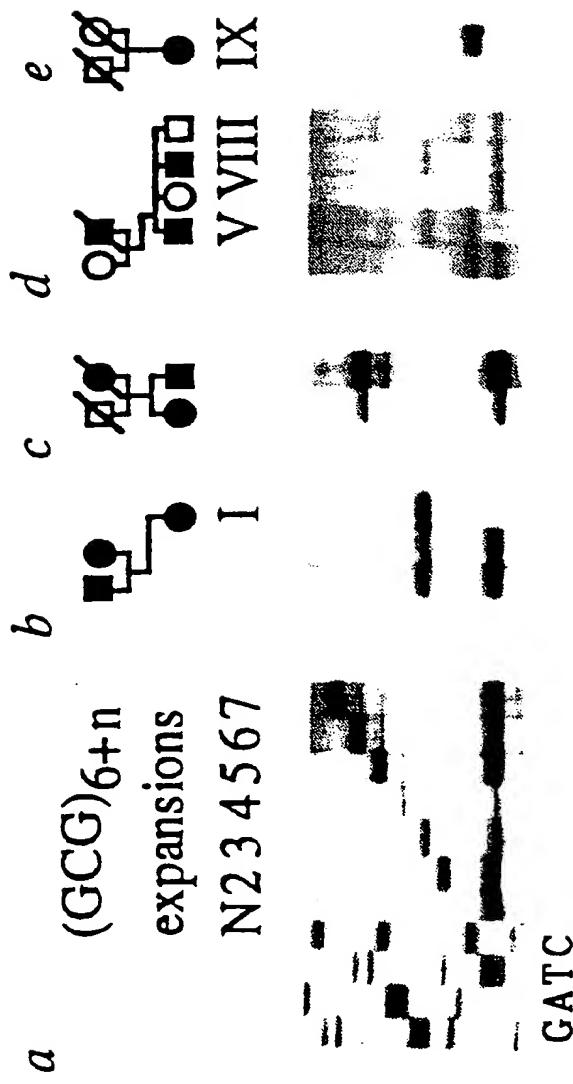
- a) designing probes and/or primers using the hGT1 gene of claim 1 and screening oculopharyngeal muscular dystrophy patients samples with said probes and/or primers; and
- b) evaluating the identified gene role in oculopharyngeal muscular dystrophy patients.

*Eco*RI restriction map (350 kb)



Human *PABII* genomic structure (6002 bp)



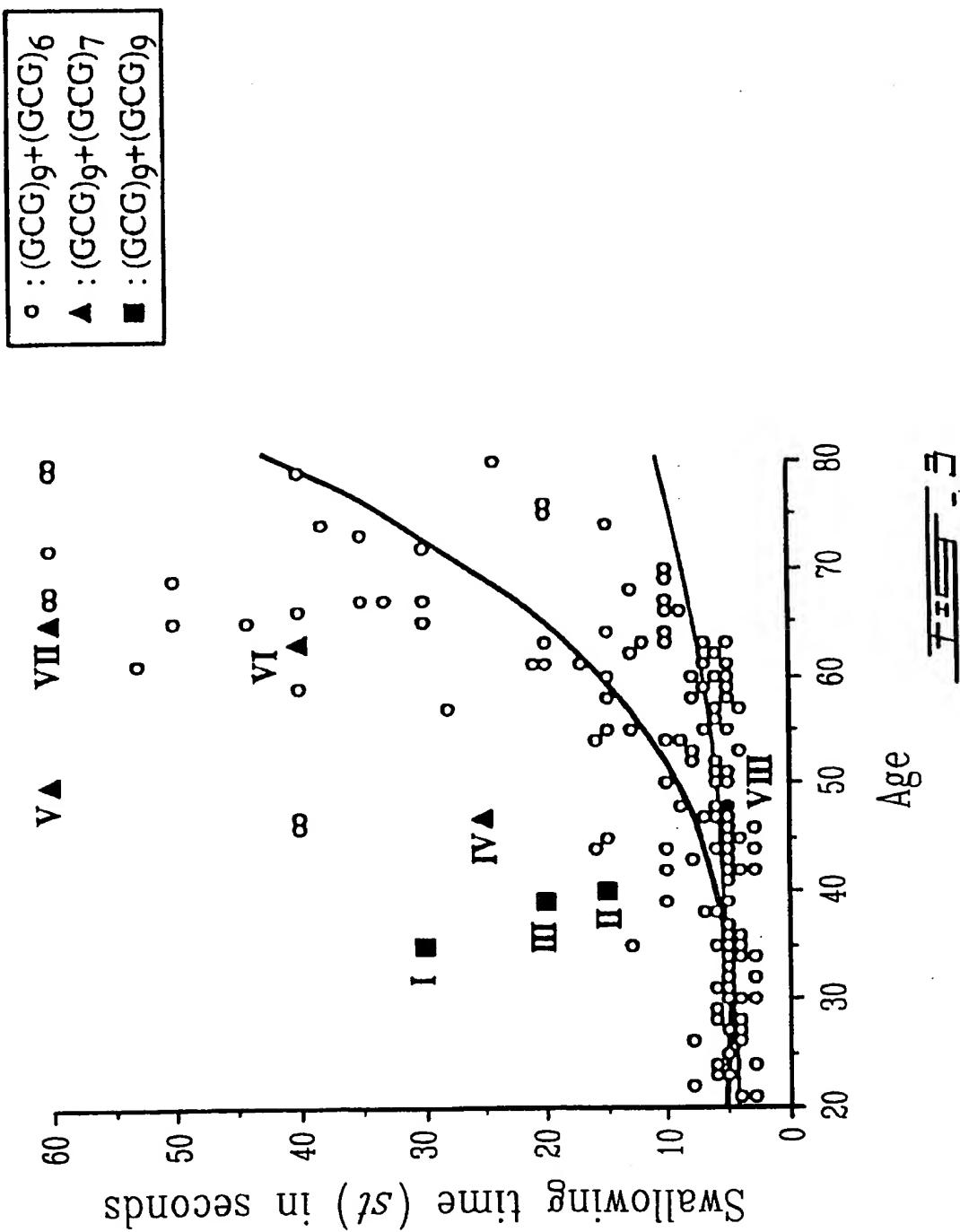


8 OPMD dominant mutations:

N: ATGGCGGGCGGGCGGGCGAGCAGCA
 ATGGCGGGCGGGCGGGCG(GCG)₂₋₇CCA

Polyalanine insertions:

N: MAAAAAA(A)2-7AAAAGAAAG
 MAAAAAA(A)2-7AAAAGAAAG



1 aatgaaggtg gacacccaaa tagccccaaat acaaattgcct gttcaaatcaa ccaaaacatct
61 aaggcagcaca tctatgttgt agcatattgc cagggcgtga gactgcgaat ataaataagga
121 accggccctc atctgcaggc gctcacaaacc tagttagcaa acagtaaaaac aattaaggcgc
181 gccgtggaca tagccctact tgccctggaa aatggggaa agctgggtt tgcaagtggtt
241 tgattgaagg gggactacat gtttagggca cagactgggt gcaggtacac ccaaaggaaac
301 gagaaggatg gaaggaaaca acatccacaa agtaaccaca agtaaccaca tgctggcgt
361 tgatttacgg ttttagact ttacctcgcc agccaaagggg ggccagtcg ttagcggtgc
421 agattggagg ggtgacattg gaaggtgtcc agggaaaaaaa aaatggaaact gggaggcaga
481 aggctacgc aaggggggg gggagggctt gacagacagg acttgtgact agtagctctg gactgaggaa
541 tcctccctgc ttcttgggtc gggagggacta gtggatgtat gtgcacaataa cctggatggg
601 gaaagtaagg tccctcctgg aatgtttcat tcacaaccc cattttcagc aacatccat
661 ctacttggtgc ttccctgggtcg agataacaatg ttccctgaaac tgctgctcg ttttggcct
721 caccggcca acagctcaat agctggcaag cagtagtatac aagatggcgg ccccttagga
781 ctggctagtcc atgtgaccctc gggtttccca agtttgaagg aatcggtcc tttcgggggc
841 aaggttcaccc tgtcacgaaa cgagtgtcac cccttcgact ctcgcaagcc aatcggtcc
901 tgaggactggg ccactgggtt gaggcgatcg gaagattgggt ctttccagt cgccctagcta
961 gggccaaatca cggaggcgtcc catacttcgc gggccccc gtagggccgg gagaaggcagg
1021 aatatcgtaa caggcggtgg gtattattac ctaaggactc gataggaggt gggacggcgt
1081 ttgatttgaca ggcagattc cctaccggaa tttagaaatt tggccagtg cccgccttag
1141 aggtggcgctt atttgattgc caagtaatat tccccaaatgg agtacttagt catggtagc
4 / 8

~~7-8~~ - 4A

1201 ggcaggcagg ttgagcttaat gagtccctccg tggccggccgc agctctccac atgcccggcg
 1261 gcgggccccca gtctgaggcg cgatggccgg ggcggccggc cggcggccggc
 1321 tgcggccgt cggggctccg gggggccat cggggccat cttgtgccc
 1381 ggaggccgg gaggggcccc cggggactac gggaaacggcc tggagttctga
 1441 ggaactggag cctgaggagg tgcgtctggaa gcccggagccg gaggcccgagg
 1501 gccggccgg ccccgccccc cccggggagg tccggccct gggcctgggt
 1561 cggcagccaa gaggaggagg ccgggtgaggaa aggaggccggg actggtcgag
 1621 cattggggac ccgggtgaggaa aggaggccgaa cggagccaggc
 1681 gaggcccaga gctcggcga ggggtggcag ggggggggtg
 5 - 1741 tggctgggc ggtctggcc ggggatgggt cggcgtatcac
 8 1801 tgattcgggc gtcacgggtg cctagtgttg ttctagagag ggttagctt
 1861 gaccctcgca tggggcgagg gaaatggccg agcatggctg
 1921 caggccacag cccctggtt ggttcctt aagctgtcct
 1981 ttaggagctg gaagctatca aagctcgagt
 2041 aaaggagcta cagaacggagg tagagaaagca
 2101 tgagtaactg gggttggcac gcggggcccc
 2161 atggggaaatg tggggataga tactcgac
 2221 tgtggccggg tcatagtcgg ttgtgtgttc
 2281 ttgggtgggtg cctcccttgg tagccttgg
 2341 tagtctacgt ctatcttct ttggtagagg
 5 - 8

2401 ataggttttc ctccaaattgg agacgcttta ggattctaag agaaagcaag ctggaaagggg
 2461 tttccccc ttcaattttaga aatgtggagt ctcaagccac ttaattttgc tcactcttaa
 2521 aaggcatttca accaaaggcca ttcatattgg atttgatttg gagggcaggaa gggattcccta
 2581 tactgttttta agtgtgtatt aattttttca atttattcgaa ttatttagtgg agtaaacctgc
 2641 tatgcacttag gcactattt cggcttgggtt gtagcaggcagg gaacaggcaca gacaaaaatc
 2701 ttggcccttca ctgagcttat gggatagtgc tggtgggtggaa agtgcaacat attggtcaag
 2761 tagaaaaacaa gtgtgtggtt ttgttaaaaaa attattttt cctgatagct ggcccgggtga
 2821 tcattgtccat tgaggagaag atggaggctg atggccggttc catctatgtt ggc当地gtga
 2881 cgtactgggg ctctgactggg ggttggggc aagttttctt tttggggat tatttaatag
 6 2941 tcctgaaagg aacatccgg gatagatgtg gtttgggtg tggggggat gtggaaaggaa
 ~ 3001 ggttaaagggt aatggaaatga tcagtaatca gcaagggtc tgggttggaa agggaaaagg
 ~ 3061 attaattccct caaattacca gatttcatgt gctttgggtgt atgatggccc agaccaaagg
 3121 ctcgggggg ttcttttgg acaggaattt gcttgggtcc ttctcccttc
 3181 atcagggtggc ctatggtgc acaggcagg agtggaaagc tcactttcat ggctgtggtt
 3241 cagtcaaccg tgttaccata ctgtgtgaca aatttagtgg ccattccaaa ggttaaagttaa
 3301 agggaggtaa gttgagataa tttaaattac agtgtacaaa tagataaatt atgtttata
 3361 ttggaggcgtt agttattttgg tggtaacaca ggtgatctgt gtcatttaag atcatggcat
 3421 taatgttgat atatcaggag ttgcaccaa atgtcttcag aggccagata aaaaaaatga
 3481 aggcttagatg tgggtggat tacgaacttag aagggggggg gcagcttcta cttggcctat
 3541 tatggcatat gggaaattcag gcccctgtgtg tcttattttt acaaatttca aqagtagct

3601 ggaaattttaa aattttaaat gatttgcgaat gatttgcgaat ttccatattg aagaatttttg
3661 acaaataaaa aatataactg cattgttagc caaaacgaag catggctgca ggttgaattt
3721 gaccgtgttag gtatttttag cctcaggatc aatacaatgac aattttttc aggtttgcgt
3781 atatagagt ctcagacaaa gagtcaatgtga ggacttccctt ggccttagat ggcctccat
3841 ttagaggaaag gcaaatacg gtaaggccat gtccatattgt gttctatgt tgataaaact
3901 ctccaggttt ctttaaggc tatcatttttgc tcatctctga ctcagggtat cccaaaaacgaa
3961 accaacagac caggcatcag cacaacagac cggggtttcc cacgagccg ctaccggcc
4021 cggaccacca actacaacag ctcccgctct cgattctaca gtggtttaa cagcaggccc
4081 cgggggtcgcc tctacaggcc aggatagatg ggctgtctccctt cttttccccc
-4141 gccccgtatg cttcctcc tctgggtctga gaaacctccc tccccccac
-4201 gtcttcaggatc actttgtctc ctggctgtgc aggttggagga aggttagttgc
-4261 agaaggcaggc ctcatcatc ttctgcagt aaaaattttgt gataagggtc
4321 cttggttcaa agaggctcc acccccagcc tttttttctt tggagttgg tggcatttg
4381 aggtgtttgc ggacaaaact ggacaaaaca gggcctccaggcactgttg
4441 gacattttgtt acttttcg gagttaggaa ggattgtaaacttgc
4501 ataccaggagg ctagctagg gatcctccca acaggcttgc
4561 attctttatc tgccaggatc ttgcaaggatc aacttctcac
4621 ggttttggcc ttgttctact tctgtctca catttaataaata
4681 ctggctttt ctaagctt acctggctat cccaggatg taggaggat ctatttgtga
4741 aggcccttagg gttaaaaac tgtggaggac tgaaaaactg gataaaaagg gggtcctttt

~~7-11~~ - 41

4801 ccttgcccct gtcttcact cagatggct tcttttcgc cacttttgg caaagtttc
 4861 tgtaagccc ccctccctt ttaggacact gccccaggttc tcccagggtgc gttactattt ctgggatcat
 4921 gggtcgggtt ttaggacact tgaacacttc ttcccttcac agtaactggg
 4981 acggggcct acggggggg gcttgtactg aactatctag tgatcacgtt aacacctaac
 5041 tctcctctt tctccaggg gccgggttag agcgacatca tggtattttcccttactaaa
 5101 aaagtgtgt a ttagggagg aagaggaaaaaaa aaaaaaaaatg agaaggaaaaaaa
 5161 taaaaaaaaaa aaaaaaaaaaa acagaagatg accttggatgg aaaaaaaaaata tttttaaaaa
 5221 aaaagatata ctgtggagg gggaggaaatc ccataactaa ctgctggagg gggacctgt
 5281 ttggggatgt gggaaaggcc caggaggatgg ggcagggggc tgcttattca ctctgggat
 5341 tcgcccatttca cagttctcaa ctggcaagg tgcttggcca tgcttccctg ccccccac
 ~5401 cccctggc ctgctcaagg ctgggtggc gtgggtggta ggagggtttt ttttacccag
 5461 ggctctggaa ggacaccaa ctgttctgtct tgtttacccctc cctccctcgcc
 5521 tttcacatc ccctcctggc tgctcctgtc cagccaggtc taccacccac cccacccctc
 5581 ttctccggc tccctggcc tccagattgc ctgggtgatct atttgttccctt cttttgtgtt
 5641 tcttttctg ttttgagggt ttttctgtc aggtttctgt agccggaaaga tctccgttcc
 5701 gctcccaagg gctccagggt aattccctt tccccccttgg gaaatgcact acttgttt
 5761 ggggggttta ggggtttt tggttttctg ttgttttctg tttttttcc
 5821 ttggccctt ttcccttta ttggggggaa atggggggaa gtgggaacag ggagggtgggaa
 5881 ggtggattt gtttatttt ttagctcatt tccagggtg ggaattttt ttttaatatgt
 5941 gtcataataa aagttgttt tgaaaaataaa aaaaaaaaaaaa
 6001 aa

~~FIGURE~~ - 4E

SEQUENCE LISTING

<110> MCGILL UNIVERSITY
ROULEAU, Guy A.
BRAIS, Bernard

<120> SHORT GCG EXPANSIONS IN THE PAG II GENE
FOR OCULOPHARYNGEAL MUSCULAR DYSTROPHY AND DIAGNOSTIC
THEREOF

<130> 1770-199PCT FC/1d

<150> CA 2,218,199
<151> 1997-12-09

<160> 6

<170> FastSEQ for Windows Version 3.0

<210> 1
<211> 57
<212> DNA
<213> Artificial Sequence

<400> 1
atggcggcgg cggcggcggc ggcagcagca atggcggcgg cggcggcggc ggcggca 57

<210> 2
<211> 35
<212> DNA
<213> Artificial Sequence

<400> 2
maaaaaaaaaa agaaggrgsm aaaaaaaaaa agaag 35

<210> 3
<211> 6002
<212> DNA
<213> Artificial Sequence

<400> 3
aatgaaggtg gacacccaaa tagcccaat acaaatgcct gttcaatcaa ccaaacatct 60
aagcagcaca tctatgtggt agcatattgc caggccgtga gactgcgaat ataaatagga 120
accggccctc atctgcaggc gctcacaacc tagttagcaa acagtaaaac aattaagcgc 180
gccgtggaca taggcccact tgtcctggga aatgagggga agctggggtt tgcagtggtt 240
tgattgaagg gggactacat gtttagaggc cagactgggt gcaggtacac ccaaagaac 300
gagaagagtg gaaggaaaca acatccacaa agtaaccaca tgctggcgta tcgaaggccg 360
tgatttacgg ttttgagact ttacctcgcc agcaaagggg ggccagtcg ttagcgtgc 420
agattggagg ggtgacattg gaagctgtcc agaaaaaaga aatggaact ggggagcaga 480
aggcctacgc aagagggcgg gacagacagg acttgtgact agtagctctg gactgaggaa 540
tcctccctgc tttctggcgc gggagagcta gtggatgatg gtgccaataa cctggatgg 600
gaaagtaagc tccctcctgg aatgcttcat tcacaacctc catttcagc aacatccat 660
ctactggcgc ttccctggcgc agatacaagt ttccctgaaac tgctgctctg ttttggcct 720
caccggcca acagctcaact agctggcaag cagtagtatac aagatggcgg ccccttagga 780
ctggctagtc atgtgaccc tc gggttccca agtttgaagc ccggcagtc ttccggggc 840
aaggttccacc tgcacgaaa cgagtgtcac cccttcgact ctcgcaagcc aatcggcattc 900
tgagactggg ccactgcggt gaggcgatcg gaagatttgtt cctttccagt cgcctagcta 960

gggccaatca	cggagcgtcc	cataacttcgc	gggcccggcc	gtaggccggg	gagaaggcagg	1020
aatatcgta	cagcgtggcg	gtattattac	ctaaggactc	gataggaggt	gggacgcgtg	1080
ttgattgaca	ggcagattc	cctaccggga	tttgcgaaatt	tggcgactg	ccgcgccttag	1140
aggtgcgtt	atttgcattgc	caagtaatat	tccccaatgg	agtactagct	catggtgacg	1200
ggcaggcagc	ttgagcta	gagtccctcg	tggccggcgc	agctctccac	atgcccggcg	1260
gcggggcccc	gtctgagcgg	cgatggcggc	ggcggcggcg	gcccgcagcag	cagcgggggc	1320
tgcgggcgg	cggggctccg	ggccggggcg	gcccgcgc	cttgcgtcccc	ggccgggtgg	1380
ggaggccggg	gagggggccc	cggggggcgc	aggggactac	ggaaacggcc	tggagctctga	1440
ggaactggag	cctgaggagc	tgctgtcgga	gcccgcgc	gagcccgagc	ccgaagagga	1500
gcccggccgg	ccccgcgc	ccccgggagc	tccgggcct	gggcctgggt	cgggagcccc	1560
cggcagccaa	gaggaggagg	aggagccggg	actggtcag	ggtgacccgg	gggacggcgc	1620
cattgaggac	ccggtgagga	aggaggcga	gcccgcgc	cgccggctgg	cgcgtcaactg	1680
gaggcccaga	gctcgccgca	gcccgcgc	gccccgggt	gggttggcg	ggaaataacg	1740
tggctggggc	gggtcggggc	ggggatgggt	cagcgtacac	tacaaggggc	cgcactggct	1800
tgatcgggg	gtcacgggt	cctagtgtt	ttctagagag	ggtagcttt	cttttatcac	1860
gaccctcgca	tggggcgagg	gaaatggccg	agcatggctg	aggcgcgc	tggccgagag	1920
cagggcacag	cccctgcgtt	ggttccttct	aagctgtct	ccataccctc	cccactata	1980
tttaggagctg	gaagctatca	aagctcgagt	cagggagatg	gagaagaag	ctgagaagct	2040
aaaggagcta	cagaacgagg	tagagaagca	gatgaatatg	agtccacctc	caggcaatgc	2100
tgagtaactg	gcgggtgcac	gcggagcccg	ggttctcggt	ttggaagggt	tgtggggagg	2160
atggggaaatg	tggggttttaga	tactcgac	cctggagctg	cttgcgtctgag	ctattatgac	2220
tgtgccgcgg	tcatagtccg	ttgtgtgtt	ctctgacctt	tgtgaggcag	aactgatatt	2280
ttgggttgtt	tagccttgc	cctcccttt	tcctgttata	attgtgttgc	tctttattct	2340
tagtctacgt	ctatcttct	ttggtagagg	ttgcgtgctc	gcatttgacc	ttcaaata	2400
atagttttc	ctccaatttg	agacgctta	ggattctaag	agaaagcaag	ctggaagggg	2460
tttccccctt	aaattctaga	aatgtggagt	ctcagccac	ttaattttgc	tcactctaa	2520
aagcatttca	accaaagcca	ttcatttaggg	atttgatttgc	gagggcagga	gggattctca	2580
tacttttta	agtgtgtatt	aattcttca	atttacgaa	ttattttagt	agtaacactgc	2640
tatgcactag	gcactattct	cggcttgcgg	gtacagcagg	gaacagcaca	gaccaaaatc	2700
tttgccttca	ctgagcttat	gggatgtgc	tgggttgtt	agtgcaacat	attggcaag	2760
tagaaaacaa	gtgtgttgtt	ttttaaaaaa	attattttt	cctgatagct	ggcccggtga	2820
tcatgtccat	tgaggagaag	atggaggctg	atgcccgtc	cattctatgtt	ggcaatgtga	2880
cgtactgggg	ctctgactgg	ggttgggggc	aagttctct	tttggggaaat	tatthaatag	2940
tcctgaaaga	acatctccgg	gatagatgt	gtttttgggt	tggaggggagt	gtgggaagga	3000
ggttaaaggt	aatggaaatga	tcataatca	gcaaaaggctc	tgggttttgg	agaaaaagag	3060
attaatttcc	caaatttacca	gatttcatgt	gctttgggt	atgtggccc	agaccaaagg	3120
ctcgggaggg	ttcttttgag	acaggaattt	gcctggtgc	tgtaaaattt	ttctcccttc	3180
atcaggttga	ctatggtgc	acagcagaag	agctggaaagc	tcactttcat	ggctgttgtt	3240
cagtcacccg	tgttaccata	ctgtgtgaca	aatttagtgg	ccatcccaaa	ggtaaagtaa	3300
agggaggtaa	gttgagataa	tttaaattac	agtgtacaaa	tagataaaatt	atgtttata	3360
ttgagcagta	agttatttgg	tgttaacaca	ggtgatctgt	gtcatttaag	atcatggcat	3420
taatgttgat	atatcaggag	ttgcacctaa	atgtcttcag	aggccagata	acaaaaatga	3480
aggctagatg	tgggtggat	tacgaactag	aaggggaggg	gcagcttcta	cttggcttat	3540
tatgcataat	ggaaatttca	gccctgtgt	tcttattttt	acaaatttca	aagtagtagct	3600
ggaattttta	aaatttaat	gatttcaat	gattgaaatt	ttccattttag	aagaattttg	3660
acaaataaaaa	aatataactg	catttgcgtt	caaaacgaa	catgcctgca	ggttgaattt	3720
gacctgttag	gtatttgtaa	cctcagagag	atacaatgc	aatttttttgc	aggtttgcgt	3780
atataagatgtt	ctcagacaaa	gagtcgttgc	ggacttcctt	ggcctttagat	gagtccctat	3840
ttagaggaag	gcaaatacg	gtaagcctat	gtccattgt	gttctagtttgc	tgtataaaact	3900
ctccaggttg	cctttaaggc	tatcattttgt	tcattctct	tcaggtgtat	ccaaaaacga	3960
accaacagac	caggcatcag	cacaacagac	cggggttttgc	cacgagcccg	ctaccggc	4020
cggaccacca	actacaacag	ctccgcgtt	cgattctaca	gtggttttaa	cagcaggccc	4080
cggggtcgcg	tctacagatgc	aggatagatg	ggctgcttct	cttccccccg	cctcccggtga	4140
gccccgtatg	cttccctctc	tctggctgt	ggAACCTCCC	tccccccacc	cctcccccgtg	4200
gtcttcagga	actttgtctc	ctgcctgtgc	aggttgagga	aggttagttgc	aggccaggcc	4260
agaaggcagc	ctcatcatct	ttctgcagt	agaaatttgc	gataagggt	gcattccctcc	4320
cttggttcaa	agaggcttcc	accccccagcc	ttttttttct	tggagttgg	tggcatttga	4380

agggtgtttgc ggacaaaact gggaggaaca gggcctccag gaagttgaaa gcactgcttg
gacattttgtt actttttcg gagttagggg ggattgaag actgaacctc ctttggaaaga
ataccagagg ctagctagtt gatcctccca acagccttgc gggaggatt tgagatactt
attctttatt tgagccagtc ttgcaagggtt aacttctcac tgggcctagt gtggtnccca
ggtttttgc ttgcttcact tctgtctcta catttaataa gacgggttag gcatataaaac
cttggcttt cataagctct acctgcctat ccccaggaggt tagggaggat ctatttgc
aggccctagg gttaaaaaac tggaggac taaaaaaactg gataaaaaagg gggcccttt
ccttgcctt gtctctcaact cagatgcgtc tcttttcgc cactgtttgg caaagtttc
tgttaagccc ccctccccct gccccagttc tcccaggtgc gttactattt ctgggatcat
ggggtcggtt ttaggacact tgaacaactt ttttcccccc ttcccttcac agtaactggg
gcagggggct acggggaggg gcttgcgtc aactatctag tgatcacgtt aacacctaac
tctccttctt tcttccaggg gcccggctag agcgacatca tggtattttcc cttactaaaa
aaagtgtgta ttaggaggag agagaggaaa aaaagaggaa agaaggaaaa aaaaagaaat
aaaaaaaaaa aaaaagaaaa acagaagatg accttgatgg aaaaaaaaaata tttttaaaaa
aaaagatata ctgtggaaagg ggggagaatc ccataactaa ctgctgagga gggacctgt
ttggggagta ggggaaggcc cagggagtgg ggcagggggc tgcttattca ctctggggat
tcgccatggc cactgtctcaa ctgcgcagc tgcttgcctt tgttttctg ccccttcac
cccttgggc ctgctcaagg gtaggtggc gtgggtggta ggagggtttt ttttaccagg
ggctctggaa ggacacccaa ctgttctgt ttttacccctt cctccctgt ttcctcgcc
ttcacagtc ccctcctgcc tgcttcgtc cagccgggtc taccacccac cccacccctc
tttctccggc tccctgcccc tccagattgc ctggtgatct atttttttcc ttttgggtt
tctttttctg ttttgagtgt ctgttttcgc aggtttctgt agccgaaaga ttcctgtttc
gctcccgcg gctccagtgt aaattccct tccccctggg gaaatgcact accttgc
gggggggtta ggggtgttt tgggtttcag ttgttttgc tttttttttt tttttttcc
tttgcccttt tccctttta ttggaggaa atgggaggaa gtggaaacag ggagggtgg
ggtggtttttt gtttattttt ttagctcatt tccagggtg ggaattttt ttaatatgt
gtcatgaata aagttgttt tgaaaataaa aaaaaaaaaaa aaaaaaaaaaa aaaaaaaaaaa
aa

```
<210> 4
<211> 19
<212> DNA
<213> Artificial Sequence
```

<400> 4
cqcaqtqcccc cqcccttaga

19

<210> 5
<211> 24
<212> DNA
<213> Artificial Sequence

<400> 5
acaqaqtqqc qccqccgccc cggc

24

```
<210> 6
<211> 23
<212> DNA
<213> Artificial Sequence
```

<400> 6
cqcaagtqccc cqcccttaqag gtg

23

INTERNATIONAL SEARCH REPORT

International Application No

PCT/CA 98/01133

A. CLASSIFICATION OF SUBJECT MATTER
 IPC 6 C12Q1/68 A01K67/027 C12N15/12

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbol(s))

IPC 6 C12Q A01K C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category ^o	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	NEMETH A ET AL: "Isolation of genomic and cDNA clones encoding bovine poly(A) binding protein II." NUCLEIC ACIDS RESEARCH, (1995 OCT 25) 23 (20) 4034-41, XP002097093 see the whole document ---	1-5
Y	DATABASE EMBL Entry HS12206 Accession Nr. U12206, 4 March 1995 XP002097097 see abstract ---	1-5
Y	DATABASE EMBL Entry AA618589 Accession Nr. AA618589, 10 October 1997 XP002097098 see abstract ---	1-5

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

Date of mailing of the international search report

18 March 1999

01/04/1999

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Reuter, U

INTERNATIONAL SEARCH REPORT

International Application No
PCT/CA 98/01133

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	TEH B T ET AL: "Oculopharyngeal muscular dystrophy (OPMD)--report and genetic studies of an Australian kindred." CLINICAL GENETICS, (1997 JAN) 51 (1) 52-5, XP002097094 see the whole document ---	1-10
A	BRAIS B ET AL: "Using the full power of linkage analysis in 11 French Canadian families to fine map the oculopharyngeal muscular dystrophy gene." NEUROMUSCULAR DISORDERS, (1997 OCT) 7 SUPPL 1 S70-4, XP002097095 see page 70 see page 74 ---	1-10
A	RIGGINS G J ET AL: "HUMAN GENES CONTAINING POLYMORPHIC TRINUCLEOTIDE REPEATS" NATURE GENETICS, vol. 2, no. 3, November 1992, pages 186-191, XP000196593 see the whole document ---	1-10
P,X	BRAIS B ET AL: "Short GCG expansions in the PABP2 gene cause oculopharyngeal muscular dystrophy" NATURE GENETICS, (1998 FEB) 18 (2) 164-7, XP002097096 see the whole document ---	1-10
P,X	WO 98 31800 A (KRISSANSEN GEOFFERY W :AUCKLAND UNISERVICES LTD (NZ); HUMAN GENOME) 23 July 1998 see page 67 - page 68; claims 1-3 -----	1-5

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CA 98/01133

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. Claims Nos.:

because they relate to subject matter not required to be searched by this Authority, namely:

Remark: Although claim 10 is directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

2. Claims Nos.:

11
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

Claim 11 is obscure. The mentioned hGT1 gene is not mentioned in claim 1, and claim 11 lacks technical details.

3. Claims Nos.:

because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

The additional search fees were accompanied by the applicant's protest.

No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

...information on patent family members

Interr. Application No

PCT/CA 98/01133

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO 9831800	A 23-07-1998	AU 5923398	A	07-08-1998
		AU 6030398	A	07-08-1998
		WO 9831799	A	23-07-1998
		AU 5927398	A	07-08-1998
		AU 6241698	A	07-08-1998
		WO 9831818	A	23-07-1998
		WO 9831801	A	23-07-1998
		WO 9831806	A	23-07-1998